

Linear Algebra – Module 2

Engineering Mathematics In Action

Course Designed by Prof. Amrik Sen

Course Conducted by Prof. Amrik Sen and Prof. Sanjay K. Bose

Plaksha University

Matrix Factorizations and Applications

LU Factorization ($\mathbf{A} = \mathbf{LU}$ where \mathbf{A} is an $n \times n$ *square matrix*)

Consider the following set of equations:

$$\mathbf{Ax}=\mathbf{b}$$

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \dots \dots \dots (i) \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \dots \dots \dots (ii) \\ & \dots & \dots \\ & \dots & \dots \\ & \dots & \dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = & b_n \dots \dots \dots (n^{th} eq.) \end{array}$$

with \mathbf{A} ($n \times n$), \mathbf{x} ($n \times 1$) and \mathbf{b} ($n \times 1$) defined as usual

Why do LU Factorization?

Since $\mathbf{Ax}=\mathbf{b}$ implies $\mathbf{LUx}=\mathbf{b}$
 or $\mathbf{Ly}=\mathbf{b}$ with $\mathbf{y}=\mathbf{Ux}$

Knowing the factors \mathbf{L} and \mathbf{U} , we
 can first solve for \mathbf{y} in $\mathbf{Ly}=\mathbf{b}$ by
 forward substitution

We then solve for \mathbf{x} in $\mathbf{Ux}=\mathbf{y}$ by
 backward substitution

$$\begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & 0 \\ m_{21} & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ m_{n1} & m_{n2} & \cdot & \cdot & m_{n(n-1)} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix} \quad \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdot & \cdot & \cdot & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdot & \cdot & \cdot & a_{2n}^{(2)} \\ \cdot & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & \cdot & a_{(n-1)n}^{(n-1)} \\ 0 & \cdot & \cdot & \cdot & 0 & a_{nn}^{(n)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$$

\mathbf{L}
 \mathbf{y}
 \mathbf{b}
 \mathbf{U}
 \mathbf{x}
 \mathbf{y}

Use Forward Substitution to get \mathbf{y}
 starting from Row 1

Use Backward Substitution to get \mathbf{x}
 starting from Row n

Why do *LU* Factorization?

Solving n equations for n unknowns

- Gauss Elimination typically requires about $2n^3/3$ operations
- The *LU* decomposition requires $n^3/3$ operations while solving *Lux=b* requires only about n^2 operations
- Therefore the *LU* decomposition method generally requires fewer operations than the Gauss Elimination method.
- One more advantage of the *LU* decomposition method is that, given a fixed **A**, changing **b** can give us the different solutions much more efficiently than Gauss Elimination. (Why? ... Because we need to do the *LU decomposition* only once!)

When can one do LU decomposition for a matrix A ?

For a **non-singular** matrix, one can always **reorder** the rows to get a matrix for which one can do LU factorization.

The objective of the reordering is to make the diagonal coefficients non-zero. Once this is done, it would be possible to do the LU factorization.

The LU factorization is unique if we ensure that the diagonal term of the L matrix is all 1's

How to do LU decomposition

Reorder rows to ensure non-zero diagonal terms


$$\begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & 0 \\ m_{21} & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & & & 0 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ m_{n1} & m_{n2} & \cdot & \cdot & m_{n(n-1)} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \dots & \dots & u_{1n} \\ 0 & u_{22} & \dots & \dots & u_{2n} \\ 0 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & u_{nn} \end{pmatrix}$$

A **L** **U**

There are several approaches to find L and U

LU Factorization Example

Explicitly solve for the u_{ij} s and the m_{ij} s

$$\begin{pmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 6 & 2 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 \\ -7 \\ 26 \end{pmatrix} \quad A = [a_{jk}] = \begin{pmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 6 & 2 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

Solution: $x_1 = 4, x_2 = -1, x_3 = \frac{1}{2}$ shown later

$$a_{11} = 3 = u_{11} \quad a_{12} = 5 = u_{12} \quad a_{13} = 2 = u_{13}$$

$$\begin{aligned} a_{21} = 0 &= m_{21}u_{11} & a_{22} = 8 &= m_{21}u_{12} + u_{22} & a_{23} = 2 &= m_{21}u_{13} + u_{23} \\ m_{21} = 0 & & u_{22} = 8 & & u_{23} = 2 & \end{aligned}$$

$$\begin{aligned} a_{31} = 6 &= m_{31}u_{11} & a_{32} = 2 &= m_{31}u_{12} + m_{32}u_{22} & a_{33} &= m_{31}u_{13} + m_{32}u_{23} + u_{33} \\ &= m_{31}(3) & &= 2(5) + m_{32}(8) & &= 2(2) + -1(2) + u_{33} \\ m_{31} = 2 & & m_{32} = -1 & & u_{33} = 6 & \end{aligned}$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 0 & 0 & 6 \end{pmatrix}$$

$L \qquad U$

Reorder rows to ensure non-zero diagonal terms

$$\mathbf{A} \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{pmatrix} = \mathbf{L} \mathbf{U}$$



$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & 0 \\ m_{21} & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ m_{n1} & m_{n2} & \cdot & \cdot & m_{n(n-1)} & 1 \end{pmatrix} \quad \mathbf{U} = \begin{pmatrix} u_{11} & u_{12} & \dots & \dots & u_{1n} \\ 0 & u_{22} & \dots & \dots & u_{2n} \\ 0 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & u_{nn} \end{pmatrix}$$

$$u_{1k} = a_{1k} \quad k = 1, \dots, n$$

$$m_{j1} = \frac{a_{j1}}{u_{11}} \quad j = 2, \dots, n$$

$$u_{jk} = a_{jk} - \sum_{s=1}^{j-1} m_{js} u_{sk} \quad k = j, \dots, n; j \geq 2$$

$$m_{jk} = \frac{1}{u_{kk}} \left(a_{jk} - \sum_{s=1}^{k-1} m_{js} u_{sk} \right) \quad j = k+1, \dots, n; k \geq 2$$

LU Factorization Example

Following a Gauss Elimination type strategy
(from Prof. Amrik Sen's notes)

$$\begin{array}{cccc|cccc}
 1 & 0 & 0 & 0 & 1 & 1 & 0 & 3 \\
 0 & 1 & 0 & 0 & 2 & 1 & -1 & 1 \\
 0 & 0 & 1 & 0 & 3 & -1 & -1 & 2 \\
 0 & 0 & 0 & 1 & -1 & 2 & 3 & -1
 \end{array}$$

I

A



$$\begin{array}{cccc|cccc}
 1 & 0 & 0 & 0 & 1 & 1 & 0 & 3 \\
 2 & 1 & 0 & 0 & 0 & -1 & -1 & -5 \\
 3 & 0 & 1 & 0 & 0 & -4 & -1 & -7 \\
 -1 & 0 & 0 & 1 & 0 & 3 & 3 & 2
 \end{array}$$

$$\begin{array}{ll}
 E_2 - 2E_1 & m_{21} = 2 \\
 E_3 - 3E_1 & m_{31} = 3 \\
 E_4 - (-1)E_1 & m_{41} = -1
 \end{array}$$



$$\begin{array}{cccc|cccc}
 1 & 0 & 0 & 0 & 1 & 1 & 0 & 3 \\
 2 & 1 & 0 & 0 & 0 & -1 & -1 & -5 \\
 3 & 4 & 1 & 0 & 0 & 0 & 3 & 13 \\
 -1 & -3 & 0 & 1 & 0 & 0 & 0 & -13
 \end{array}$$

L

U

$$\begin{array}{ll}
 E_3 - 4E_2 & m_{32} = 4 \\
 E_4 - (-3)E_2 & m_{42} = -3
 \end{array}$$

$$m_{43} = 0$$

LU Factorization Example

Following a Gauss Elimination type strategy

$$\begin{pmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 6 & 2 & 8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 3 & 5 & 2 \\ 0 & 1 & 0 & 0 & 8 & 2 \\ 0 & 0 & 1 & 6 & 2 & 8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 3 & 5 & 2 \\ 0 & 1 & 0 & 0 & 8 & 2 \\ 0 & 0 & 1 & 6 & 2 & 8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 3 & 5 & 2 \\ 0 & 1 & 0 & 0 & 8 & 2 \\ 2 & 0 & 1 & 0 & -8 & 4 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 0 & 3 & 5 & 2 \\ 0 & 1 & 0 & 0 & 8 & 2 \\ 2 & -1 & 1 & 0 & 0 & 6 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 0 & 0 & 6 \end{pmatrix}$$

L U

Solving a set of Linear Equations using LU Factorization

$$Ax = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 0 & 0 & 6 \end{pmatrix} x = b = \begin{pmatrix} 8 \\ -7 \\ 26 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix} y = b = \begin{pmatrix} 8 \\ -7 \\ 26 \end{pmatrix} \Rightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 8 \\ -7 \\ 3 \end{pmatrix}$$

L U L y

$$Ux = y \Rightarrow \begin{pmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 0 & 0 & 6 \end{pmatrix} x = \begin{pmatrix} 8 \\ -7 \\ 3 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 1/2 \end{pmatrix}$$

Solving a set of Linear Equations using *LU* Factorization (*Prof. Amrik Sen's example*)

$$Ax = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix} x = b = \begin{pmatrix} 4 \\ 1 \\ -3 \\ 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{pmatrix} y = b = \begin{pmatrix} 4 \\ 1 \\ -3 \\ 4 \end{pmatrix} \Rightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 4 \\ -7 \\ 13 \\ -13 \end{pmatrix}$$

L
 U
 L
 y

$$Ux = y \Rightarrow \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{pmatrix} x = \begin{pmatrix} 4 \\ -7 \\ 13 \\ -13 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

Orthogonal Basis and Gram-Schmidt Orthogonalization

Two vectors \vec{u}_1 and \vec{u}_2 are **orthogonal** if and only if $\langle \vec{u}_1, \vec{u}_2 \rangle = 0$

The vectors $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m \in R^n$ are **orthonormal** if and only if $\langle \vec{u}_i, \vec{u}_j \rangle = \delta_{ij} = 1 \quad i = j$

Example: $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n \in R^n$ are orthonormal $= 0 \quad i \neq j$

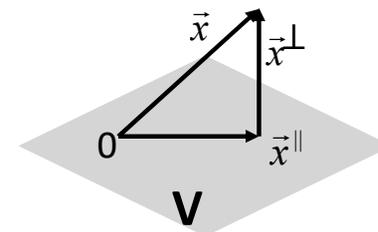
Properties of orthonormal vectors

1. Orthonormal vectors are (automatically) linearly independent
2. Orthonormal vectors $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n \in R^n$ form a basis in R^n

Orthogonal projection and orthogonal complement:

Let $\vec{x} \in \mathbb{R}^n$ and a subspace V of \mathbb{R}^n . Then we can write $\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$, where $\vec{x}^{\parallel} \in V$ and $\vec{x}^{\perp} \in V^{\perp}$. The above representation is **unique**.

Here $V^{\perp} = \{\vec{x} \in \mathbb{R}^n : \langle \vec{v}, \vec{x} \rangle = 0, \forall \vec{v} \in V\}$. The transformation $T(\vec{x}) = \text{proj}_V \vec{x} = \vec{x}^{\parallel}$ from \mathbb{R}^n to \mathbb{R}^n is linear. $V^{\perp} = \text{Ker}(T)$.



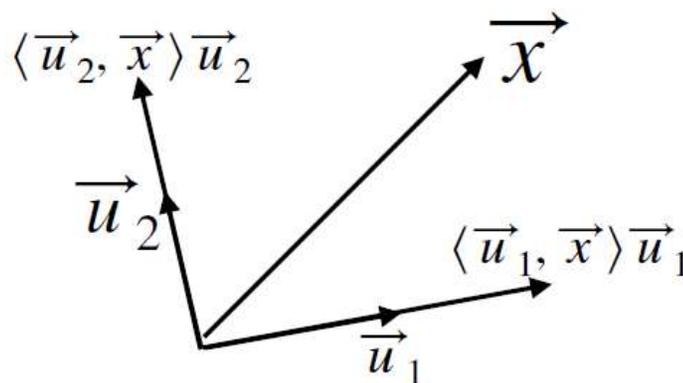
How do we compute \vec{x}^{\parallel} ?

Consider an orthonormal basis of V : $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m \in V$ which is a subspace of \mathbb{R}^n . Then

$$\vec{x}^{\parallel} = \langle \vec{u}_1, \vec{x} \rangle \vec{u}_1 + \dots + \langle \vec{u}_m, \vec{x} \rangle \vec{u}_m; \quad \forall \vec{x} \in \mathbb{R}^n.$$

Consequently, consider an orthonormal basis of \mathbb{R}^n : $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$. Then any $\vec{x} \in \mathbb{R}^n$,

$$\vec{x} = \langle \vec{u}_1, \vec{x} \rangle \vec{u}_1 + \dots + \langle \vec{u}_n, \vec{x} \rangle \vec{u}_n.$$



\vec{x}^{\perp} Component of \mathbf{x} perpendicular to V

\vec{x}^{\parallel} Projection of \mathbf{x} on V

Properties of V^\perp , the Orthogonal Complement of V :

Consider a subspace $V \in \mathbb{R}^n$. Then, we have the following important results about V^\perp , the orthogonal complement of V

1. V^\perp is a subspace of \mathbb{R}^n .
2. $V \cap V^\perp = \{0\}$.
3. $\dim(V) + \dim(V^\perp) = n$.
4. $(V^\perp)^\perp = V$.

Example:

Consider the subspace $V = \text{Im}(A)$ of \mathbb{R}^4 where $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}$. Find \vec{x}^{\parallel} for $\vec{x} = \begin{pmatrix} 1 \\ 3 \\ 1 \\ 7 \end{pmatrix}$.

Solution: Recall that the column space of A is $\text{Im}(A)$. It can be easily checked that the column vectors of A are orthogonal by taking their scalar product. Thus we can construct an orthonormal basis of $\text{Im}(A)$. The basis vectors \vec{u}_1, \vec{u}_2 are -

$$\vec{u}_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{pmatrix} \text{ and } \vec{x}^{\parallel} = \langle \vec{u}_1, \vec{x} \rangle \vec{u}_1 + \langle \vec{u}_2, \vec{x} \rangle \vec{u}_2 = 6\vec{u}_1 + 2\vec{u}_2 = \begin{pmatrix} 4 \\ 2 \\ 2 \\ 4 \end{pmatrix} \quad \text{What would be } \vec{x}^{\perp}?$$

Verify that $\vec{x} - \vec{x}^{\parallel} \perp \vec{u}_1, \vec{u}_2$ perpendicular to both \vec{u}_1 and \vec{u}_2

Why are orthonormal basis vectors useful?

1. We know that if we have some basis $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ of an n -dimensional vector space W , then any vector $\vec{x} \in W$ can be written as $\vec{x} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_n \vec{v}_n$ (i.e. as a linear combination of the basis vectors). However, there is no first-principles or convenient way of finding the unique coefficients $\alpha_1, \alpha_2, \dots, \alpha_n$ except by explicit guesswork calculations. On the contrary, if we do have an orthonormal basis set $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$ then any vector can be written as a linear combination of this orthonormal basis set as

$$\vec{x} = \beta_1 \vec{u}_1 + \beta_2 \vec{u}_2 + \dots + \beta_n \vec{u}_n$$

where the coefficients $\beta_i = \langle \vec{u}_i, \vec{x} \rangle, \forall i = 1, 2, \dots, n$ are **uniquely determined**.

2. Orthogonality guarantees linear independence.

Orthogonal Transformation: A linear transformation $T: V \rightarrow V$ in \mathbb{R}^n that preserves the inner product. For each pair u, v of elements of V , we will then have -

$$\langle u, v \rangle = \langle Tu, Tv \rangle$$

Why are orthogonal transformations useful?

1. Orthogonal transformations are **metric preserving transformations**, i.e. if

$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n \text{ is orthogonal, then } \|T(\vec{x})\| = \|\vec{x}\|, \forall \vec{x} \in \mathbb{R}^n.$$

(length preserving transformation)

(Note that if $T(\vec{x}) = A\vec{x}$ is an orthogonal transformation, then we say that A is an orthogonal matrix. For such an A , we have that $A^T A = I$ or $A^{-1} = A^T$)

2. Orthogonal transformations are **angle preserving transformations** for orthogonal vectors.

For example, if $\vec{u} \perp \vec{w}$, then $T(\vec{u}) \perp T(\vec{w})$